

OIPE

RAW SEQUENCE LISTING

DATE: 01/10/2002

PATENT APPLICATION: US/10/015,085

TIME: 14:38:07

Input Set : A:\10271037999.txt

Output Set: N:\CRF3\01102002\J015085.raw

ENTERED

4 <110> APPLICANT: Langermann, Solomon R.
 5 Hultgren, Scott J.
 6 Hung, Chia-Suei
 7 Bouckaert, Julie
 9 <120> TITLE OF INVENTION: Mutant Proteins, High Potency Inhibitory Antibodies, and
 FimCH
 10 Crystal Structure
 12 <130> FILE REFERENCE: 10271-037
 C--> 14 <140> CURRENT APPLICATION NUMBER: US/10/015,085
 C--> 15 <141> CURRENT FILING DATE: 2001-12-10
 17 <160> NUMBER OF SEQ ID NOS: 50
 19 <170> SOFTWARE: PatentIn version 3.0
 21 <210> SEQ ID NO: 1
 22 <211> LENGTH: 726
 23 <212> TYPE: DNA
 24 <213> ORGANISM: E. coli
 26 <220> FEATURE:
 27 <221> NAME/KEY: CDS
 28 <222> LOCATION: (1)...(723)
 30 <400> SEQUENCE: 1
 31 atg agt aat aaa aac gtc aat gta agg aaa tcg cag gaa ata aca ttc 48
 32 Met Ser Asn Lys Asn Val Asn Val Arg Lys Ser Gln Glu Ile Thr Phe
 33 1 5 10 15
 35 tgc ttg ctg gca ggt atc ctg atg ttc atg gca atg atg gtt gcc gga 96
 36 Cys Leu Leu Ala Gly Ile Leu Met Phe Met Ala Met Met Val Ala Gly
 37 20 25 30
 39 cgc gct gaa gcg gga gtg gcc tta ggt gcg act cgc gta att tat ccg 144
 40 Arg Ala Glu Ala Gly Val Ala Leu Gly Ala Thr Arg Val Ile Tyr Pro
 41 35 40 45
 43 gca ggg caa aaa caa gtg caa ctt gcc gtg aca aat aat gat gaa aat 192
 44 Ala Gly Gln Lys Gln Val Gln Leu Ala Val Thr Asn Asn Asp Glu Asn
 45 50 55 60
 47 agt acc tat tta att caa tca tgg gtg gaa aat gcc gat ggt gta aag 240
 48 Ser Thr Tyr Leu Ile Gln Ser Trp Val Glu Asn Ala Asp Gly Val Lys
 49 65 70 75 80
 51 gat ggt cgt ttt atc gtg acg cct cct ctg ttt gcg atg aag gga aaa 288
 52 Asp Gly Arg Phe Ile Val Thr Pro Pro Leu Phe Ala Met Lys Gly Lys
 53 85 90 95
 55 aaa gag aat acc tta cgt att ctt gat gca aca aat aac caa ttg cca 336
 56 Lys Glu Asn Thr Leu Arg Ile Leu Asp Ala Thr Asn Asn Gln Leu Pro
 57 100 105 110
 59 cag gac cgg gaa agt tta ttc tgg atg aac gtt aaa gcg att ccg tca 384
 60 Gln Asp Arg Glu Ser Leu Phe Trp Met Asn Val Lys Ala Ile Pro Ser
 61 115 120 125
 63 atg gat aaa tca aaa ttg act gag aat acg cta cag ctc gca att atc 432
 64 Met Asp Lys Ser Lys Leu Thr Glu Asn Thr Leu Gln Leu Ala Ile Ile
 65 130 135 140
 67 agc cgc att aaa ctg tac tat cgc ccg gct aaa tta gcg ttg cca ccc 480

RAW SEQUENCE LISTING

DATE: 01/10/2002

PATENT APPLICATION: US/10/015,085

TIME: 14:38:07

Input Set : A:\10271037999.txt

Output Set: N:\CRF3\01102002\J015085.raw

```

68 Ser Arg Ile Lys Leu Tyr Tyr Arg Pro Ala Lys Leu Ala Leu Pro Pro
69 145 150 155 160
71 gat cag gcc gca gaa aaa tta aga ttt cgt cgt agc gcg aat tct ctg 528
72 Asp Gln Ala Ala Glu Lys Leu Arg Phe Arg Arg Ser Ala Asn Ser Leu
73 165 170 175
75 acg ctg att aac ccg aca ccc tat tac ctg acg gta aca gag ttg aat 576
76 Thr Leu Ile Asn Pro Thr Tyr Tyr Leu Thr Val Thr Glu Leu Asn
77 180 185 190
79 gcc gga acc cgg gtt ctt gaa aat gca ttg gtg cct cca atg ggc gaa 624
80 Ala Gly Thr Arg Val Leu Glu Asn Ala Leu Val Pro Pro Met Gly Glu
81 195 200 205
83 agc acg gtt aaa ttg cct tct gat gca gga agc aat att act tac cga 672
84 Ser Thr Val Lys Leu Pro Ser Asp Ala Gly Ser Asn Ile Thr Tyr Arg
85 210 215 220
87 aca ata aat gat tat ggc gca ctt acc ccc aaa atg acg ggc gta atg 720
88 Thr Ile Asn Asp Tyr Gly Ala Leu Thr Pro Lys Met Thr Gly Val Met
89 225 230 235 240
91 gaa taa 726
92 Glu
96 <210> SEQ ID NO: 2
97 <211> LENGTH: 241
98 <212> TYPE: PRT
99 <213> ORGANISM: E. coli
101 <400> SEQUENCE: 2
102 Met Ser Asn Lys Asn Val Asn Val Arg Lys Ser Gln Glu Ile Thr Phe
103 1 5 10 15
104 Cys Leu Leu Ala Gly Ile Leu Met Phe Met Ala Met Met Val Ala Gly
105 20 25 30
106 Arg Ala Glu Ala Gly Val Ala Leu Gly Ala Thr Arg Val Ile Tyr Pro
107 35 40 45
108 Ala Gly Gln Lys Gln Val Gln Leu Ala Val Thr Asn Asn Asp Glu Asn
109 50 55 60
110 Ser Thr Tyr Leu Ile Gln Ser Trp Val Glu Asn Ala Asp Gly Val Lys
111 65 70 75 80
112 Asp Gly Arg Phe Ile Val Thr Pro Pro Leu Phe Ala Met Lys Gly Lys
113 85 90 95
114 Lys Glu Asn Thr Leu Arg Ile Leu Asp Ala Thr Asn Asn Gln Leu Pro
115 100 105 110
116 Gln Asp Arg Glu Ser Leu Phe Trp Met Asn Val Lys Ala Ile Pro Ser
117 115 120 125
118 Met Asp Lys Ser Lys Leu Thr Glu Asn Thr Leu Gln Leu Ala Ile Ile
119 130 135 140
120 Ser Arg Ile Lys Leu Tyr Tyr Arg Pro Ala Lys Leu Ala Leu Pro Pro
121 145 150 155 160
122 Asp Gln Ala Ala Glu Lys Leu Arg Phe Arg Arg Ser Ala Asn Ser Leu
123 165 170 175
124 Thr Leu Ile Asn Pro Thr Pro Tyr Tyr Leu Thr Val Thr Glu Leu Asn
125 180 185 190
126 Ala Gly Thr Arg Val Leu Glu Asn Ala Leu Val Pro Pro Met Gly Glu

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/015,085

DATE: 01/10/2002

TIME: 14:38:07

Input Set : A:\10271037999.txt

Output Set: N:\CRF3\01102002\J015085.raw

```

127          195          200          205
128 Ser Thr Val Lys Leu Pro Ser Asp Ala Gly Ser Asn Ile Thr Tyr Arg
129          210          215          220
130 Thr Ile Asn Asp Tyr Gly Ala Leu Thr Pro Lys Met Thr Gly Val Met
131 225          230          235          240
132 Glu
135 <210> SEQ ID NO: 3
136 <211> LENGTH: 903
137 <212> TYPE: DNA
138 <213> ORGANISM: E. coli
140 <220> FEATURE:
141 <221> NAME/KEY: CDS
142 <222> LOCATION: (1)...(900)
144 <400> SEQUENCE: 3
145 atg aaa cga gtt att acc ctg ttt gct gta ctg ctg atg ggc tgg tcg      48
146 Met Lys Arg Val Ile Thr Leu Phe Ala Val Leu Leu Met Gly Trp Ser
147 -20          -15          -10
149 gta aat gcc tgg tca ttc gcc tgt aaa acc gcc aat ggt acc gct atc      96
150 Val Asn Ala Trp Ser Phe Ala Cys Lys Thr Ala Asn Gly Thr Ala Ile
151 -5          -1 1          5          10
153 cct att ggc ggt ggc agc gcc aat gtt tat gta aac ctt gcg ccc gtc      144
154 Pro Ile Gly Gly Gly Ser Ala Asn Val Tyr Val Asn Leu Ala Pro Val
155          15          20          25
157 gtg aat gtg ggg caa aac ctg gtc gtg gat ctt tcg acg caa atc ttt      192
158 Val Asn Val Gly Gln Asn Leu Val Val Asp Leu Ser Thr Gln Ile Phe
159          30          35          40
161 tgc cat aac gat tat ccg gaa acc att aca gac tat gtc aca ctg caa      240
162 Cys His Asn Asp Tyr Pro Glu Thr Ile Thr Asp Tyr Val Thr Leu Gln
163          45          50          55
165 cga ggc tcg gct tat ggc ggc gtg tta tct aat ttt tcc ggg acc gta      288
166 Arg Gly Ser Ala Tyr Gly Gly Val Leu Ser Asn Phe Ser Gly Thr Val
167 60          65          70          75
169 aaa tat agt ggc agt agc tat cca ttt cct acc acc agc gaa acg ccg      336
170 Lys Tyr Ser Gly Ser Ser Tyr Pro Phe Pro Thr Thr Ser Glu Thr Pro
171          80          85          90
173 cgc gtt gtt tat aat tcg aga acg gat aag ccg tgg ccg gtg gcg ctt      384
174 Arg Val Val Tyr Asn Ser Arg Thr Asp Lys Pro Trp Pro Val Ala Leu
175          95          100          105
177 tat ttg acg cct gtg agc agt gcg ggc ggg gtg gcg att aaa gct ggc      432
178 Tyr Leu Thr Pro Val Ser Ser Ala Gly Gly Val Ala Ile Lys Ala Gly
179          110          115          120
181 tca tta att gcc gtg ctt att ttg cga cag acc aac aac tat aac agc      480
182 Ser Leu Ile Ala Val Leu Ile Leu Arg Gln Thr Asn Asn Tyr Asn Ser
183          125          130          135
186 gat gat ttc cag ttt gtg tgg aat att tac gcc aat aat gat gtg gtg      528
187 Asp Asp Phe Gln Phe Val Trp Asn Ile Tyr Ala Asn Asn Asp Val Val
188 140          145          150          155
190 gtg cct act ggc ggc tgc gat gtt tct gct cgt gat gtc acc gtt act      576
191 Val Pro Thr Gly Gly Cys Asp Val Ser Ala Arg Asp Val Thr Val Thr

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/015,085

DATE: 01/10/2002

TIME: 14:38:07

Input Set : A:\10271037999.txt

Output Set: N:\CRF3\01102002\J015085.raw

```

192          160          165          170
194 ctg ccg gac tac cct ggt tca gtg cca att cct ctt acc gtt tat tgt      624
195 Leu Pro Asp Tyr Pro Gly Ser Val Pro Ile Pro Leu Thr Val Tyr Cys
196          175          180          185
198 gcg aaa agc caa aac ctg ggg tat tac ctc tcc ggc aca acc gca gat      672
199 Ala Lys Ser Gln Asn Leu Gly Tyr Tyr Leu Ser Gly Thr Thr Ala Asp
200          190          195          200
202 gcg ggc aac tcg att ttc acc aat acc gcg tcg ttt tca cct gca cag      720
203 Ala Gly Asn Ser Ile Phe Thr Asn Thr Ala Ser Phe Ser Pro Ala Gln
204          205          210          215
206 ggc gtc ggc gta cag ttg acg cgc aac ggt acg att att cca gcg aat      768
207 Gly Val Gly Val Gln Leu Thr Arg Asn Gly Thr Ile Ile Pro Ala Asn
208 220          225          230          235
210 aac acg gta tcg tta gga gca gta ggg act tcg gcg gtg agt ctg gga      816
211 Asn Thr Val Ser Leu Gly Ala Val Gly Thr Ser Ala Val Ser Leu Gly
212          240          245          250
214 tta acg gca aat tat gca cgt acc gga ggg cag gtg act gca ggg aat      864
215 Leu Thr Ala Asn Tyr Ala Arg Thr Gly Gly Gln Val Thr Ala Gly Asn
216          255          260          265
218 gtg caa tcg att att ggc gtg act ttt gtt tat caa taa      903
219 Val Gln Ser Ile Ile Gly Val Thr Phe Val Tyr Gln
220          270          275
224 <210> SEQ ID NO: 4
225 <211> LENGTH: 300
226 <212> TYPE: PRT
227 <213> ORGANISM: E. coli
229 <400> SEQUENCE: 4
230 Met Lys Arg Val Ile Thr Leu Phe Ala Val Leu Leu Met Gly Trp Ser
231 -20          -15          -10
232 Val Asn Ala Trp Ser Phe Ala Cys Lys Thr Ala Asn Gly Thr Ala Ile
233 -5          -1 1          5          10
234 Pro Ile Gly Gly Gly Ser Ala Asn Val Tyr Val Asn Leu Ala Pro Val
235 15          20          25
236 Val Asn Val Gly Gln Asn Leu Val Val Asp Leu Ser Thr Gln Ile Phe
237 30          35          40
238 Cys His Asn Asp Tyr Pro Glu Thr Ile Thr Asp Tyr Val Thr Leu Gln
239 45          50          55
240 Arg Gly Ser Ala Tyr Gly Gly Val Leu Ser Asn Phe Ser Gly Thr Val
241 60          65          70          75
242 Lys Tyr Ser Gly Ser Ser Tyr Pro Phe Pro Thr Thr Ser Glu Thr Pro
243 80          85          90
244 Arg Val Val Tyr Asn Ser Arg Thr Asp Lys Pro Trp Pro Val Ala Leu
245 95          100          105
246 Tyr Leu Thr Pro Val Ser Ser Ala Gly Gly Val Ala Ile Lys Ala Gly
247 110          115          120
248 Ser Leu Ile Ala Val Leu Ile Leu Arg Gln Thr Asn Asn Tyr Asn Ser
249 125          130          135
250 Asp Asp Phe Gln Phe Val Trp Asn Ile Tyr Ala Asn Asn Asp Val Val
251 140          145          150          155

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/015,085

DATE: 01/10/2002

TIME: 14:38:07

Input Set : A:\10271037999.txt

Output Set: N:\CRF3\01102002\J015085.raw

```

252 Val Pro Thr Gly Gly Cys Asp Val Ser Ala Arg Asp Val Thr Val Thr
253                               160                               165                               170
254 Leu Pro Asp Tyr Pro Gly Ser Val Pro Ile Pro Leu Thr Val Tyr Cys
255                               175                               180                               185
256 Ala Lys Ser Gln Asn Leu Gly Tyr Tyr Leu Ser Gly Thr Thr Ala Asp
257                               190                               195                               200
258 Ala Gly Asn Ser Ile Phe Thr Asn Thr Ala Ser Phe Ser Pro Ala Gln
259                               205                               210                               215
260 Gly Val Gly Val Gln Leu Thr Arg Asn Gly Thr Ile Ile Pro Ala Asn
261 220                               225                               230                               235
262 Asn Thr Val Ser Leu Gly Ala Val Gly Thr Ser Ala Val Ser Leu Gly
263                               240                               245                               250
264 Leu Thr Ala Asn Tyr Ala Arg Thr Gly Gly Gln Val Thr Ala Gly Asn
265                               255                               260                               265
266 Val Gln Ser Ile Ile Gly Val Thr Phe Val Tyr Gln
267                               270                               275
269 <210> SEQ ID NO: 5
270 <211> LENGTH: 30
271 <212> TYPE: DNA
C--> 272 <213> ORGANISM: Artificial
274 <220> FEATURE:
275 <223> OTHER INFORMATION: Description of Artificial Sequence: Primer
277 <400> SEQUENCE: 5
278 ggggggaatt caccgcgagg gatgattgta 30
280 <210> SEQ ID NO: 6
281 <211> LENGTH: 28
282 <212> TYPE: DNA
C--> 283 <213> ORGANISM: Artificial
285 <220> FEATURE:
286 <223> OTHER INFORMATION: Description of Artificial Sequence: Primer
288 <400> SEQUENCE: 6
289 ccagtaggca ccaccacatc attattgg 28
291 <210> SEQ ID NO: 7
292 <211> LENGTH: 48
293 <212> TYPE: DNA
C--> 294 <213> ORGANISM: Artificial
296 <220> FEATURE:
297 <223> OTHER INFORMATION: Description of Artificial Sequence: Primer
299 <400> SEQUENCE: 7
300 ctggtcggta aatgcctggt cagcggcctg taaaaccgcc aatggtac 48
302 <210> SEQ ID NO: 8
303 <211> LENGTH: 48
304 <212> TYPE: DNA
C--> 305 <213> ORGANISM: Artificial
307 <220> FEATURE:
308 <223> OTHER INFORMATION: Description of Artificial Sequence: Primer
310 <400> SEQUENCE: 8
311 gtaccattgg cgggtttaca ggccgctgac caggcattta ccgaccag 48
313 <210> SEQ ID NO: 9

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/015,085

DATE: 01/10/2002

TIME: 14:38:08

Input Set : A:\10271037999.txt

Output Set: N:\CRF3\01102002\J015085.raw

L:14 M:270 C: Current Application Number differs, Replaced Application Number
L:15 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:272 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:5
L:283 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:6
L:294 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:7
L:305 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:8
L:316 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:9
L:327 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:10
L:338 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:11
L:349 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:12
L:360 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:13
L:371 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:14
L:382 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:15
L:393 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:16
L:404 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:17
L:415 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:18
L:426 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:19
L:437 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:20
L:448 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:21
L:459 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:22
L:470 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:23
L:481 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:24
L:492 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:25
L:503 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:26
L:514 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:27
L:525 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:28
L:536 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:29
L:547 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:30
L:558 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:31
L:569 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:32
L:580 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:33
L:591 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:34
L:602 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:35
L:613 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:36
L:624 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:37
L:635 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:38
L:646 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:39
L:657 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:40
L:668 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:41
L:679 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:42
L:690 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:43
L:701 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:44
L:712 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:45
L:723 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:46
L:734 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:47
L:745 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:48
L:756 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:49
L:767 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:50